Sampler algorithm for non-convex inverse problem

Pierre Palud

PhD directed by
Pierre Chainais, Franck Le Petit

with the collaboration of
Emeric Bron, Pierre-Antoine Thouvenin

Ecole Centrale de Lille, CRIStAL, LERMA

financed by CNRS via 80|Prime

The Galaxy volume

Solar systems fill $\sim 3 \cdot 10^{-10}$ of the volume of the galaxy

~~Most of the galaxy: empty!~~

Most of the galaxy: **Interstellar Medium**!

Observations of GMC: Orion B in visible frequencies

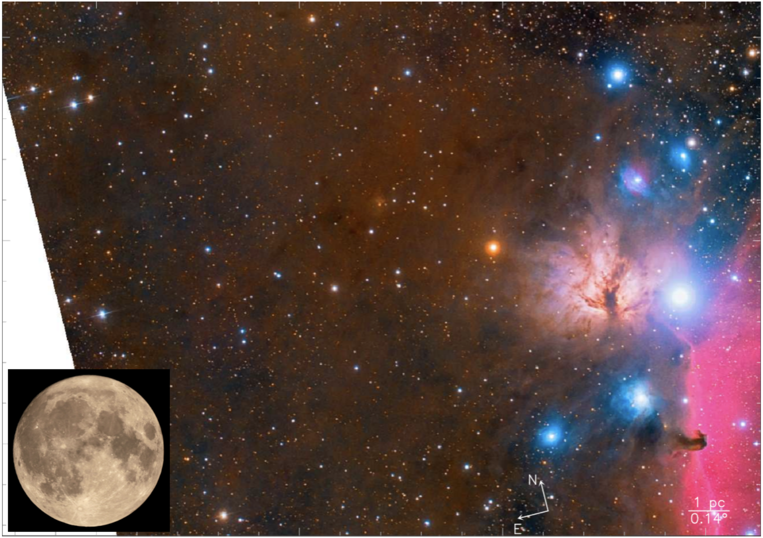


Figure: Image from Pety et al. [2016]

Observations of GMC: Orion B in visible frequencies

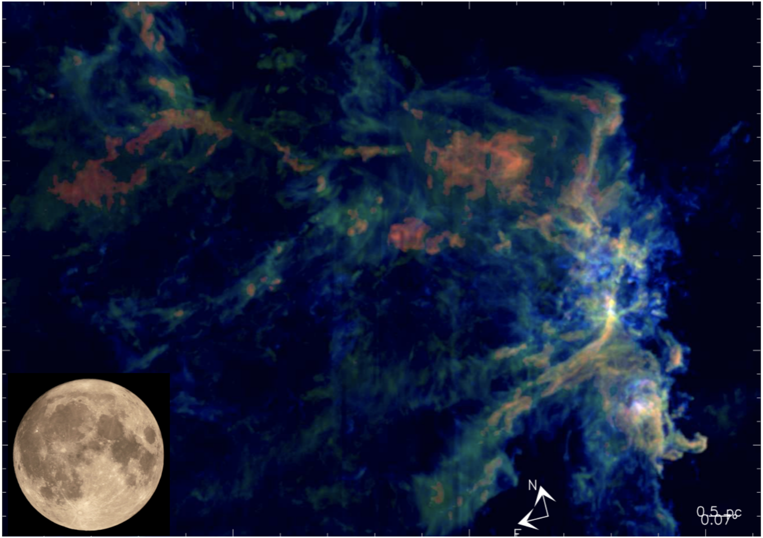


Figure: Image from Pety et al. [2016]
blue: ^{12}CO, green: ^{13}CO, red: C^{18}O

Photo-Dissociation Region (PDR)

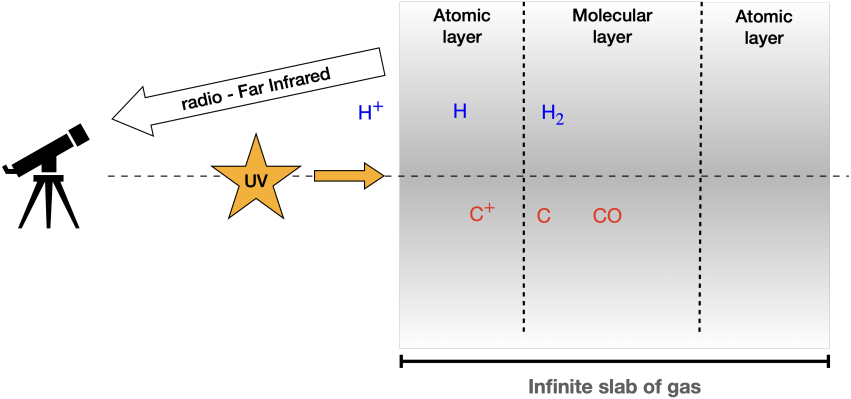


Figure: Structure of a PDR

Photo-Dissociation Region (PDR)

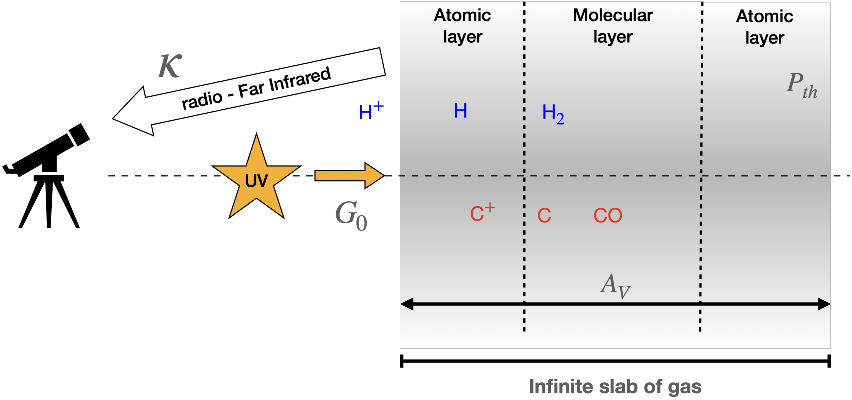
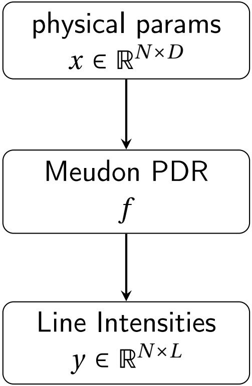


Figure: Structure of a PDR

Meudon PDR code: numerical simulation of a PDR

- Introduced in Le Petit et al. [2006].
- for stationary 1D slab of gas, solves:
  1. radiative transfer
  2. chemistry
  3. thermal balance

  that are all **coupled**!



Can we infer x from y and f?
no ground truth → with **credibility intervals**

Current state of the art in astrophysics

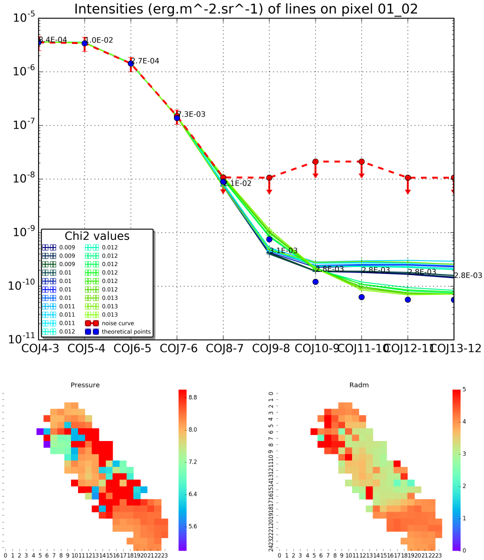


Figure: MLE maps inference

Bayesian map inversion

inference with credibility interval

$\Downarrow$

a posteriori probability distribution $\mathbb{P}\left[x \mid y\right]$

$$\underbrace{\mathbb{P}\left[x \mid y\right]}_{\textbf{a posteriori}} \propto \underbrace{\mathbb{P}\left[y \mid x\right]}_{\text{likelihood}} \times \underbrace{\mathbb{P}\left[x\right]}_{\text{a priori}}$$

Complex distribution

$\Longrightarrow$ impossible to manipulate as is

$\Longrightarrow$ sampling with MCMC

Observation model

$$\forall n, \ell,\ y_{n,\ell} = \max\left\{\omega,\ \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell}\right\}$$

with

- $\epsilon^{(a)}_{n,\ell}$: additive noise (thermal, instruments)
- $\epsilon^{(m)}_{n,\ell}$: multiplicative noise (calibration error)
- ω : minimum detectable value by telescope

Observation model

$$\forall n, \ell,\ y_{n,\ell} = \max\left\{\omega,\ \epsilon^{(m)}_{n,\ell} f_\ell(x_n) + \epsilon^{(a)}_{n,\ell}\right\}$$

with

- $\epsilon^{(a)}_{n,\ell}$: additive noise (thermal, instruments)
- $\epsilon^{(m)}_{n,\ell}$: multiplicative noise (calibration error)
- ω : minimum detectable value by telescope

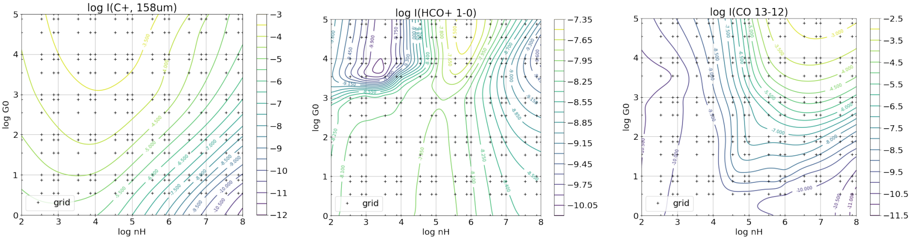


Figure: f_ℓ for some lines ℓ, for 1 pixel

Including a priori information

a priori information on x:

- spatial regularization: L2 penalty on image Laplacian
- f_ℓ: estimated from a grid
  → **constraint of belonging to a cube** (convex enveloppe of grid)
  ⚠non smooth prior ⚠but can be tempered with a smooth penalty

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior
classic MCMC algorithm (e.g., MALA) OK

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior

~~classic MCMC algorithm (e.g., MALA) OK~~ because of two difficulties

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior

~~classic MCMC algorithm (e.g., MALA) OK~~ because of two difficulties

- f and ∇f cover **various decades**
  $\implies$ classic methods **inefficient**

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior

~~classic MCMC algorithm (e.g., MALA) OK~~ because of two difficulties

- f and ∇f cover **various decades**
  $\implies$ classic methods **inefficient**

- **non-convex** negative log-posterior (because of Meudon PDR code)
  $\implies$ need to avoid being **trapped in local minima**

Sampler

smooth prior + **smooth** likelihood $\implies$ **smooth** posterior

~~classic MCMC algorithm (e.g., MALA) OK~~ because of two difficulties

- f and ∇f cover **various decades**
  $\implies$ classic methods **inefficient**
  ✓ **Preconditioned MALA kernel with RMSProp**

- **non-convex** negative log-posterior (because of Meudon PDR code)
  $\implies$ need to avoid being **trapped in local minima**
  ✓ **Multiple-Try Metropolis (MTM) kernel**

Sampler

smooth prior + **smooth** likelihood $\Longrightarrow$ **smooth** posterior

~~classic MCMC algorithm (e.g., MALA) OK~~ because of two difficulties

- f and ∇f cover **various decades**
  $\Longrightarrow$ classic methods **inefficient**
  ✓ **Preconditioned MALA kernel with RMSProp**

- **non-convex** negative log-posterior (because of Meudon PDR code)
  $\Longrightarrow$ need to avoid being **trapped in local minima**
  ✓ **Multiple-Try Metropolis (MTM) kernel**

Final sampler : random combination of these two kernels

Illustration: Gaussian Mixture in a square

Illustration that our algorithm explores interesting **local minima** :

- mixture of 20 gaussians
- with constraint $x \in [-10, 10] \times [-10, 10]$

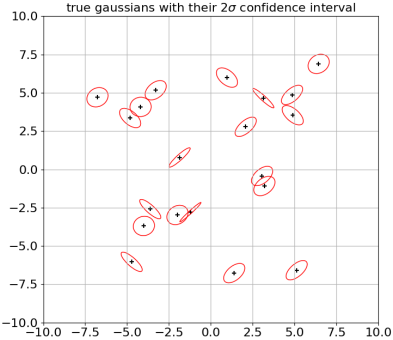
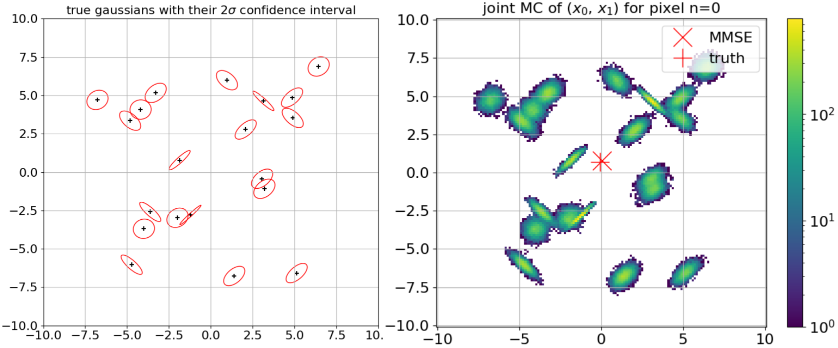


Illustration: Gaussian Mixture in a square

Illustration that our algorithm explores interesting **local minima** :

- mixture of 20 gaussians
- with constraint $x \in [-10, 10] \times [-10, 10]$



Toy case 1: Time Series Inversion

$$y_{n,\ell} = \epsilon^{(m)}_{n,\ell} f(x_n) + \epsilon^{(a)}_{n,\ell} \text{ with } f : x \in \mathbb{R} \mapsto e^x,\ \sigma_a = 1,\ \sigma_m \sim 10\%$$

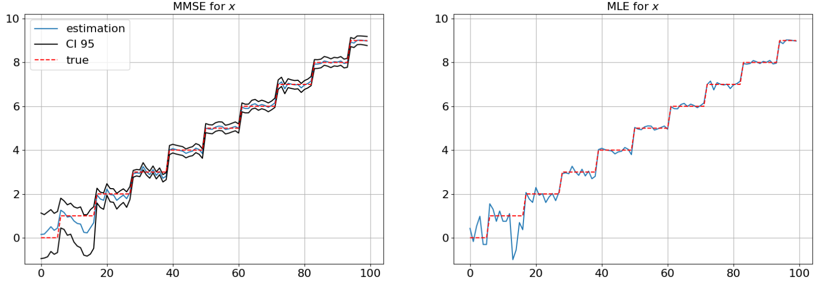
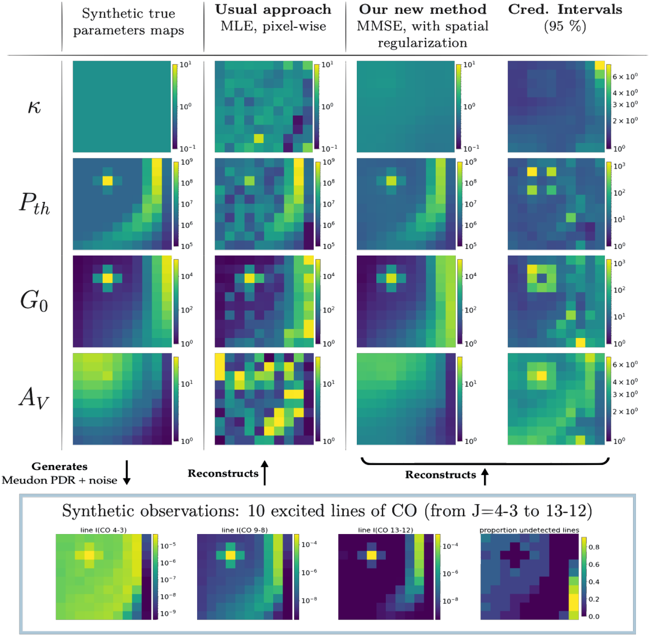


Table: Estimation Summary

estimator	MSE	SNR
MMSE	3.6	28.8
MLE	10.5	24.1

Astrophysical Toy case: Map Inversion



Application to NGC 7023 (1 pixel)

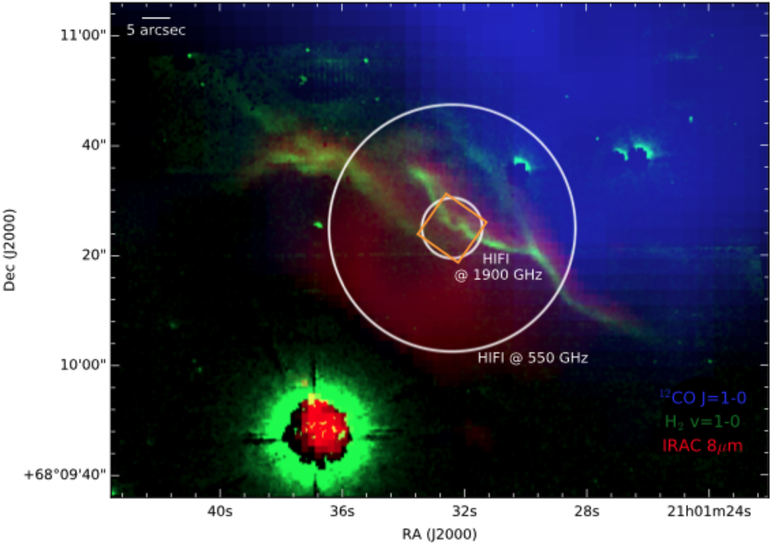
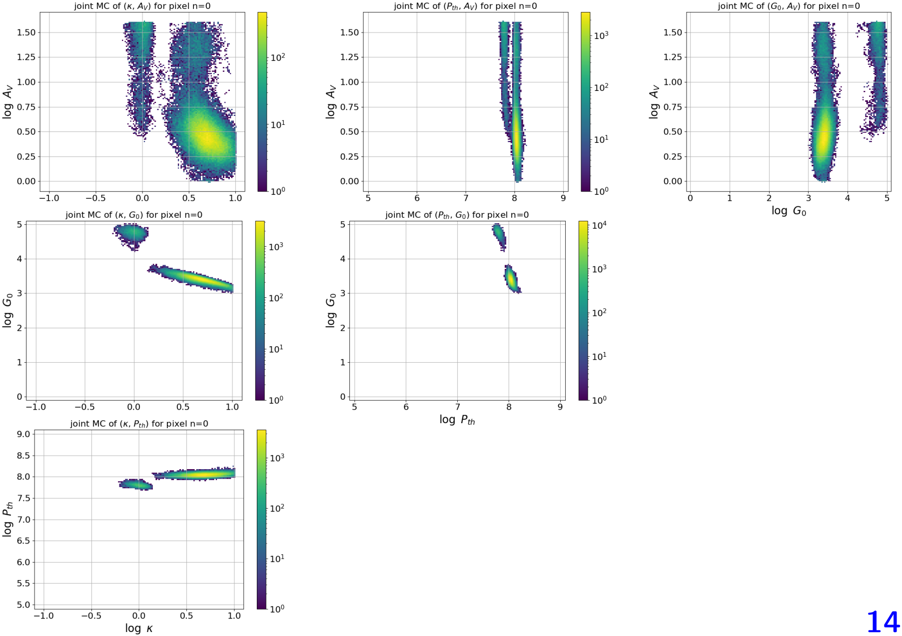


Figure: from Joblin et al. [2018]

Application to NGC 7023 (1 pixel)



Summary

- Definition of a MCMC sampler with
  1. P-MALA kernel to tackle regularity issues
  2. MTM kernel to tackle the non-log-concavity of the posterior
- Evaluation of the method on toy data
- Application to real world data

Thank you for your attention!

C. Joblin, E. Bron, C. Pinto, P. Pilleri, F. Le Petit, M. Gerin, J. Le Bourlot, A. Fuente, O. Berne, J. R. Goicoechea, E. Habart, M. Köhler, D. Teyssier, Z. Nagy, J. Montillaud, C. Vastel, J. Cernicharo, M. Röllig, V. Ossenkopf-Okada, and E. A. Bergin. Structure of photodissociation fronts in star-forming regions revealed by *Herschel* observations of high-J CO emission lines. Astronomy & Astrophysics, 615:A129, July 2018. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201832611. URL https://www.aanda.org/10.1051/0004-6361/201832611.

F. Le Petit, C. Nehme, J. Le Bourlot, and E. Roueff. A Model for Atomic and Molecular Interstellar Gas: The Meudon PDR Code. The Astrophysical Journal Supplement Series, 164(2):506–529, June 2006. ISSN 0067-0049, 1538-4365. doi: 10.1086/503252. URL https://iopscience.iop.org/article/10.1086/503252.

J. Pety, V. Guzman, J. Orkisz, H. Liszt, M. Gerin, E. Bron, S. Bardeau, J. Goicoechea, P. Gratier, F. Petit, F. Levrier, K. Oberg, E. Roueff, and A. Sievers. The anatomy of the Orion B Giant Molecular Cloud: A local template for studies of nearby galaxies. Astronomy & Astrophysics, 599, Nov. 2016. doi: 10.1051/0004-6361/201629862.